UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
"Language is Spatial": Experimental Evidence for Image Schemas of Concrete and Abstract Verbs

Permalink
https://escholarship.org/uc/item/9vs820bx

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 23(23)

ISSN
1069-7977

Authors
Richardson, Daniel C.
Spivey, Michael J.
Edelman, Shimon
et al.

Publication Date
2001

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

"Language is Spatial": Experimental Evidence for Image Schemas of Concrete and Abstract Verbs

Daniel C. Richardson (dcr18@cornell.edu),
Michael J. Spivey (spivey@cornell.edu)
Shimon Edelman (se37@cornell.edu)
Adam J. Naples (ajn23@cornell.edu)
Department of Psychology, Cornell University
Ithaca, NY 14853 USA

Abstract

Cognitive linguistics and experimental psychology have produced tantalizing hints that a substantial portion of language is encoded in the mind in the form of spatial representations that are grounded in perception and action. Researchers represent these spatial aspects using "image schemas" that depict verbs of motion or spatial prepositions via a 2-D layout of generic icons. In two experiments, we tested naïve subjects' intuitions about such image schemas for concrete action verbs as well as abstract action verbs and psychological predicates. A substantial agreement across subjects was observed in both a forced choice task and a free form computer-based drawing task, for both concrete verbs and abstract verbs. In addition to providing support for the generality of image schemas, the data provide a set of norms for future online studies of spatial representations underlying real-time language processing.

Introduction

Many theorists have argued for a spatial component to language. The arguments are commonly set against an amodal view of representation which defines items in some formal symbolic system. The motivations for proposing an alternative to the symbolic approach range from difficulties in implementing a symbolic system (Barsalou, 1999), commonalties between 'parsing' in the visual system and in language (Landau & Jackendoff, 1993), capturing subtle asymmetries and nuances of linguistic representation in a spatial, schematic way (Langacker, 1987; Talmy, 1983), and a more general account of the mind as an embodied, experiential system (Lakoff, 1987).

If we accept the idea that there is a spatial or perceptual basis to the representation of linguistic items, it would be reasonable to assume that there is some commonality between these representations across different speakers, since by and large we communicate successfully. Therefore, we might expect that there would be a consensus among subjects when we ask them to draw simple diagrams representing words. Theorists such as Langacker (1987) have produced large bodies of diagrammatic linguistic representations, arguing that they are constrained by linguistic observations and intuitions in the same way that 'well formedness' judgements inform more traditional linguistic theories. However, it remains to be seen whether naïve subjects share these intuitions and forms of representation. Therefore, in the same way that psycholinguists use norming studies to support claims of preference for certain grammatical structures, we propose to survey a large number of subjects and see if there is a consensus amongst their spatial representations of words.

Recent work has also documented the mapping between spatial linguistic terms and the mental representation of space (eg Hayward & Tarr, 1995; Carlson-Radvansky, Covey & Lattanzi, 1999; Schober, 1995). Although there are consistencies in the ways in which spatial language is produced and comprehended (eg Hayward & Tarr, 1995), the exact mapping appears to be modulated by such factors as visual context (Spivey-Knowlton, Tanenhaus, Eberhard & Sedivy, 1998), the common ground between conversants (Schober, 1995) and the functional attributes of the objects being described (Carlson-Radvansky et al., 1999).

When language refers directly to explicit spatial properties, locations, and relationships in the world, it is quite natural to expect those linguistic representations to have at least some degree of overlap in their format. Spatial language terms appear to be grounded, at least somewhat, in perceptual (rather than amodal) formats of representation. However, an important component of the work presented herein involves testing for this representational format in an arena of language that does *not* exhibit any literal spatial properties: abstract verbs (such as 'respect' and 'succeed'). Much work in cognitive linguistics has in fact argued that many linguistic and conceptual representations (even abstract ones) are based on metaphorical connections to spatially laid out "image schemas" (Gibbs, 1996; Lakoff, 1987; Langacker, 1987; Talmy, 1983). This work suggests that if consistency across subjects is observed for spatial depictions of *concrete* verbs, then one should also expect such consistency for *abstract* verbs. Experimental evidence for this kind of broad consensus among speakers would extend the "language is spatial" hypothesis beyond spatial terms, and make some experimentally supported, albeit preliminary, claims about abstract language as well.

There are various old and new results suggesting that there is some consistency among speakers in the *visual imagery* associated with certain ideas and concepts. For example, Scheerer and Lyons (1957) asked subjects to match the referents ‘gold’, ‘silver’, and ‘iron’ with three drawings which had previously been produced by other naive subjects. At least one set of these drawings (which resembled sine, saw tooth, and square waves, respectively), were correctly matched by 85% of the subjects. Lakoff (1987) offers anecdotal evidence that when asked to describe their image of an idiom such as 'keeping at arms length' people have a considerable degree of commonality in their responses, including details such as the angle of the protagonist's hand. Similarly, Gibbs, Strom and Spivey-Knowlton (1997) carried out empirical work querying subjects about their mental images of proverbs such as 'a rolling stone gathers no moss' and found a surprising degree of agreement – even about fine details such as the stone bouncing slightly as it rolled. Experimental work has shown that the listing the features of a concept involves something akin to visually inspecting its properties (cf Barsalou, 1999).

This approach extends beyond the simple visual properties of a concept, towards more schematic or spatial representations. Barsalou’s (1999) perceptual symbol system theory endorses the view held by several theorists (e.g. Lakoff, 1987; Gibbs, 1996) that to some degree abstract concepts are represented by a metaphoric relation to more concrete domains. For example, it is argued that the concept of ‘anger’ draws on a concrete representation of ‘liquid in a container under pressure’. There is some debate over how central these metaphorical aspects are to the representation of abstract concepts. But for the representation of time, at least, there is strong experimental evidence that subjects’ reasoning is structured by a metaphorical relation to space.

Boroditsky (1999) observed that English speakers tend to use horizontal spatial metaphors when talking about time, whereas Mandarin speakers use both horizontal and vertical. In a reaction time study, speakers from both languages were asked true/false questions about time (e.g. ‘March comes earlier than April’) It was found that Mandarin speakers responded faster when they had been presented with vertical rather than horizontal spatial primes, and the reverse was true for English speakers. The result is particularly impressive since both groups carried out the experiment in English. Boroditsky (2000) identified two schemas that are used in both spatial and temporal domains: ego moving and time/object moving. She found that the use of one type of schema in the spatial domain, would prime a judgement of the same schema in a temporal domain.

In this paper, we empirically test the claim that between subjects there is a coherence to the imagistic aspects of their linguistic representations. To this end we will address two questions – *Do subjects agree with each other about the spatial component of different verbs?* And, *Across a forced choice and an open ended response task, are the same spatial representations being accessed?* It would be of further interest if the subjects’ diagrams bore resemblance to those proposed by theorists such as Langacker (1987). However, as with more standard norming studies, the real value of the data will be in allowing us to generate prototypical representations that could be used as stimuli for studies of online language comprehension.

Experiment 1

Methods

Subjects 173 Cornell undergraduates participated in exchange for course credit.

Design We selected 30 verbs to fill out a concreteness by spatial layout, 2x3 factor design. Using the MRC psycholinguistic database, we divided the words into high and low concreteness. These two concreteness groups were each divided into 3 groups based on the expected primary axes of their image schemas (vertical, horizontal, and neutral), based on our survey of the cognitive grammar literature. This 2x3 factor design was filled with a list of 30 verbs. Each was placed in the past tense in the form of a simple rebus sentence, with circle and square symbols representing agents and patients.

The subjects were presented with a single page, containing a list of the verbs and four pictures, labelled A to D. Each one contained a circle and a square aligned along a vertical or horizontal axis, connected by an arrow pointing up, down, left or right. Since we didn't expect any interesting item variation between left or right placement of the circle or square, the horizontal schemas differed only in the direction of the arrow.

For each sentence, subjects were asked to select one of the four sparse images that best depicted the event described by the sentence (Figure 1)

The items were randomised in three different orders, and crossed with two different orderings of the images. The six lists were then distributed randomly to subjects.

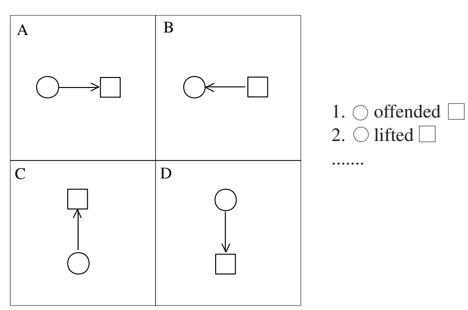

Figure 1: Example of the questionnaire in Experiment 1.

Results

Subjects’ responses are summarised in Table 2. The most frequently chosen image column is in bold for each verb. On average, for any given verb, the particular image orientation that was most popular was chosen by 63% of the subjects. The second most popular was chosen by 21%, the third by

10% and the fourth by 5%. This suggests a substantial degree of agreement between subjects.

To test our predictions concerning the primary axes of the verbs' image schemas, we converted the forced choice data into axis angles. The left and right image schemas were assigned an angle of 0, and the up and down image schemas a value of 90. See Table 2.

A two-way ANOVA by-items analysis revealed a significant main effect of expected axis ($F(2,24)=30.30$, $p<0.0001$), and the effect of concreteness did not approach significance ($F(1,24)=1.84$, $p>0.18$). There was, however, a significant interaction ($F(2,24)=5.28$, $p<0.02$), indicating that the effect of expected axis was more dramatic for concrete verbs than for abstract verbs. Planned comparisons revealed that, even among the abstract verbs, the mean axis angle of the expected-horizontal verbs was lower than that of the neutral verbs and the mean axis angle of the vertical verbs was greater than that of the neutral verbs (all *p*s<.05).

Table 1: Percentage of subjects choosing each image

Concreteness	Expected Axis	Verb	Up	Down	Left	Right
HIGH	Horizontal	fled	7.2	4.2	**80.8**	7.8
		pointed at	7.2	3.6	0	**89.2**
		pulled	6	5.4	**75.4**	13.2
		pushed	7.2	3.6	1.2	**88**
		walked	9	3.6	24	**62.9**
	Neutral	hunted	9.6	20.4	1.8	**68.3**
		impacted	7.2	37.1	3	**52.7**
		perched	12	**76**	6.6	5.4
		showed	15	9	10.2	**65.9**
		smashed	3.6	**66.5**	1.2	28.7
	Vertical	bombed	4.8	**86.8**	1.8	6.6
		flew	37.7	**44.3**	15	3
		floated	32.9	**56.3**	7.8	3
		lifted	**87.4**	9.6	2.4	0.6
		sank	22.2	**71.9**	4.2	1.8
LOW	Horizontal	argued with	11.4	13.8	12.6	**62.3**
		gave to	8.4	9.6	1.2	**80.8**
		offended	9	31.7	24.6	**34.7**
		rushed	10.2	10.8	23.4	**55.1**
		warned	10.8	22.2	6	**61.1**
	Neutral	owned	5.4	**55.7**	18.6	20.4
		regretted	19.8	24	**41.3**	15
		rested	14.4	36.5	**40.1**	9
		tempted	16.8	11.4	**45.5**	26.3
		wanted	15.6	7.8	15.6	**61.1**
	Vertical	hoped	**45.5**	15.6	7.2	31.7
		increased	**73.7**	7.2	9.6	9
		obeyed	22.8	4.2	**64.7**	8.4
		respected	**53.9**	3	14.4	28.7
		succeeded	**40.1**	35.9	10.8	13.2
		Means	20.9	26.2	19	33.8

Table 2: Mean Axis angle.

Expected Axis / Concreteness	Horizontal	Neutral	Vertical
High	10	46	82
Low	25	37	55

Discussion

It appears that there is a considerable degree of agreement between subjects. This consistency was seen in both concrete verbs of motion, eg 'lifted', and abstract verbs, such as 'respected'. Yet it could be argued that this coherence mainly reflects the artificial and limited nature of the forced choice ask, rather than a commonality of deeper significance between subjects' representations. In our next experiment, we allowed subjects to create their own image schemas in an open response task.

Experiment 2

In this experiment, we asked subjects to create their own representation of the sentences using a simple computer based drawing environment.

Method

Subjects Twenty-four Cornell University undergraduates participated in exchange for course credit. None of these subjects had participated in Experiment 1.

Design Subjects were presented at random with a sentence from Experiment 1. They were given as much time as they required to draw a schematic representation of the sentence. When they had finished, they clicked a done button and were given the next sentence.

The drawing environment is shown in Figure 2. Subjects could drag the shapes on to central canvas. Any number of shapes could be used, and they could be repositioned. Subjects could also use up to 3 arrows. By holding down modifier keys, the arrows could be re-sized and rotated.

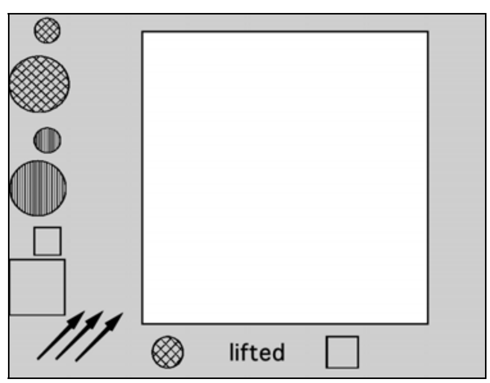

Figure 2: Screen shot from Experiment 2.

Results

Subjects spent approximately a minute completing each drawing. Some subjects produced quite sparse, schematic representations; others attempted more complex depictions. Figures 3 and 4 show a random selection of drawings of the concrete verb ‘argued with’ and the more abstract verb ‘respected’.

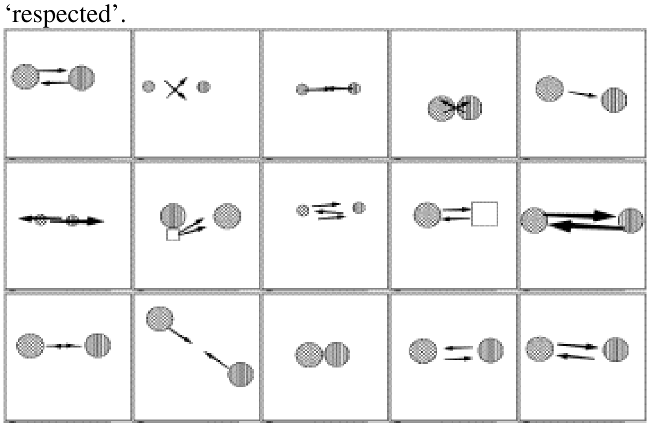

Figure 3: Example depictions of “ARGUED WITH”.

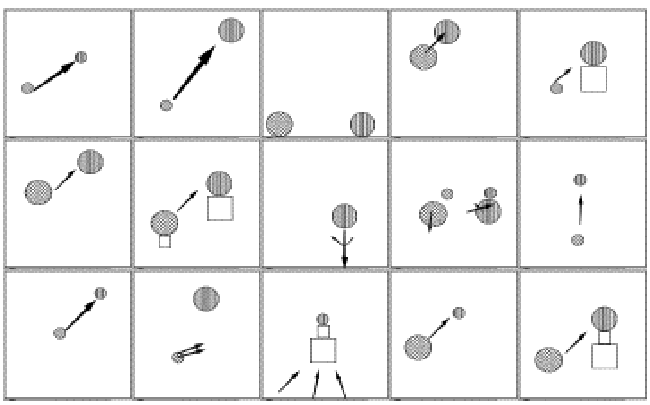

Figure 4: Example depictions of “RESPECTED”.

The majority of subjects appeared to represent the verbs schematically using quite sparse images. However, there were a few subjects who, despite the limitations of the drawing toolbox, attempted to *pictorially* represent the verbs. For example, in the third and fourth figures in the second row of Figure 4, we can see that the subjects have drawn humanoid figures, using the arrows as arms. Indeed, since they were the only items that could be rotated and resized, the arrows were often used as generic lines forming a pictorial drawing. For this reason, we decided to ignore the arrows in our analysis, and focus on the relative positions of objects.

Using the coordinates of objects in the drawing, we defined the ‘aspect angle’, as a value between 0 and 90 which reflects the horizontal versus vertical extent of each drawing. If one imagines a box drawn around the center points of all objects in a picture, the aspect angle is the angle of a diagonal line connecting the lower-left and upper-right corners of the box. If the objects are aligned on a horizontal axis, the aspect angle would be 0; on a vertical axis, 90.

Note that the aspect angle collapses left-right and top-bottom mirror reflections of a drawing. We decided to use this measure since we were primarily interested in the horizontal versus vertical aspect of each drawing. In addition, the initial starting orientation of the arrows (Figure 2) might bias subject towards a right rather than left, and an upwards rather than downwards layout in their drawings: this bias would be avoided in calculating the aspect angle. Figure 5 graphically represents the aspect angle data in what we have termed a ‘radar’ plot. Each verb’s mean aspect angle (solid line) is shown together with its standard error (shaded fan area), and included is the mean axis angle of that verb in the forced choice task of Experiment 1 (dashed line). The means for each condition are shown in the final column of Figure 5.

These results were subjected to a ANOVA items analysis. The only significant effect was of expected axis ($F(2,24)=6.69, p<0.005$). The mean aspect angle for the horizontal group was 21˚, neutral 36˚ and vertical 45˚.

Discussion

Despite the free form nature of the task, it seems that there was a reasonably high degree of agreement between subjects. The mean standard error for all verbs was 6.5 degrees. Previous work has found that subjects consistently place the flow of action from left to right when depicting events (Chatterjee, 2001), but our subjects employed contrasting horizontal and vertical image schemas as well. Moreover, there is considerable consistency between the drawings and the results of the forced choice task. We observed a significant correlation between the mean aspect angles for the verbs in the two tasks ($R= 0.71$).

General Discussion

We have presented data that suggest there is an impressive degree of coherence in the spatial, schematic components of some linguistic representations. Two different tasks attempted to tap these representations. A forced choice task with very sparse images appeared to produce comparable results to a creative, open response task. In this sense, the data suggest a positive answer to two of our opening questions - Do subjects agree with each other? and Do the two tasks assess the same underlying representations?

We would also argue that our third question - *Do naive subjects agree with trained linguistic intuitions?* – can be given a qualified ‘yes’. In both experiments, the expected axis had a significant effect on the orientation of subjects’ responses. Figure 5 reveals some informative cases where our expectations were defeated by subjects. For example, in our neutral condition, both ‘perched’ and ‘rested’ were consistently given a vertical aspect angle by subjects in both tasks. This observation highlights the importance of using normative methodologies to accompany traditional linguistic methodologies.

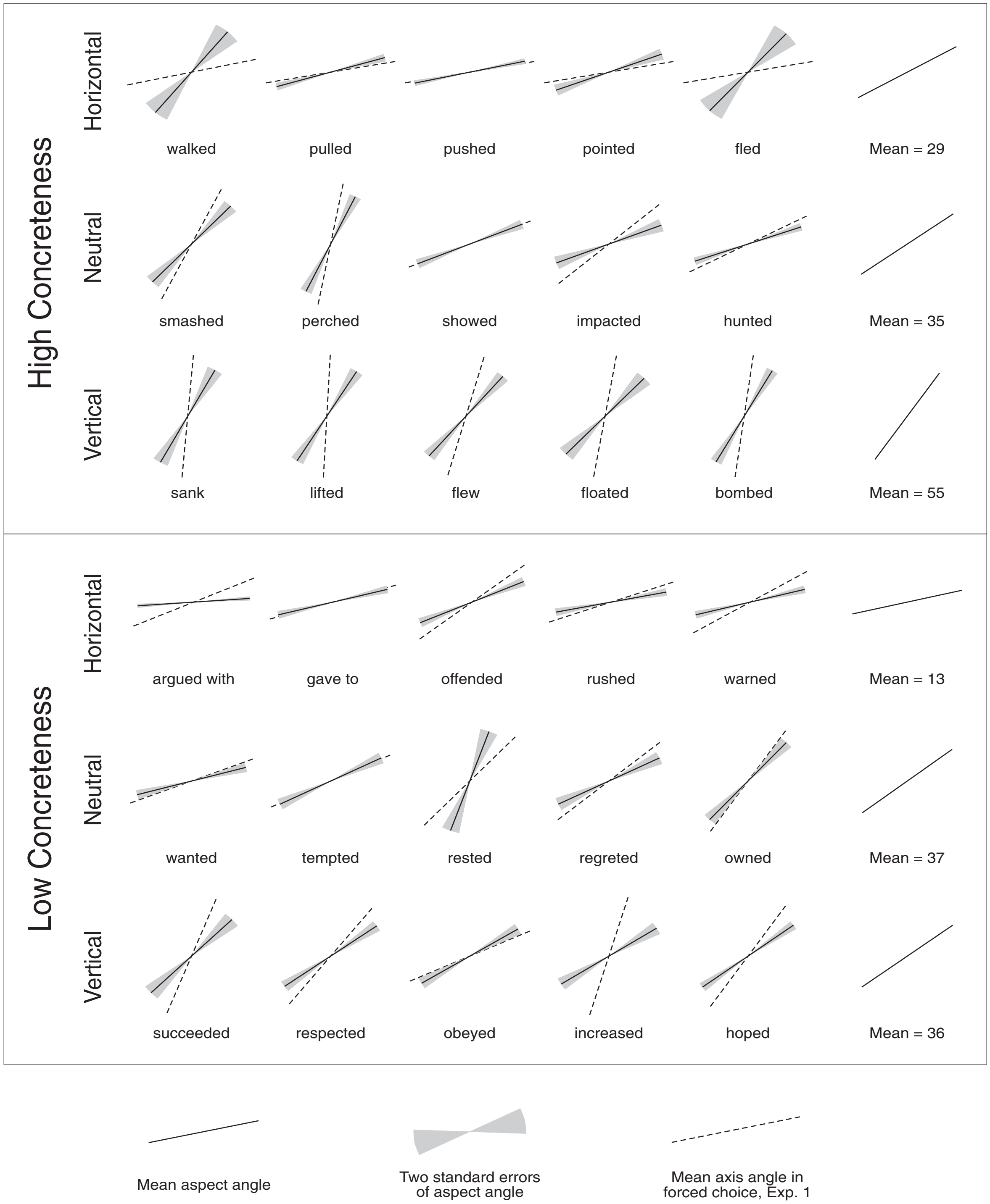

Figure 5: 'Radar' plots of mean aspect angles in subjects' drawings, Exp.2

Although these norming studies demonstrate considerable (and perhaps surprising) agreement between the intuitions of cognitive linguists and naïve subjects, we would argue that the true value of these results is in the predictions they generate for real-time language processing. Just as offline word similarity ratings predict online performance in word priming tasks, we hope that our offline data will predict effects of spatial priming for online language comprehension. In this way, we can further test whether the spatial formats of linguistic representation suggested by these results are indeed fundamental components of language processing in natural situations, and not just artefacts of contemplative metalinguistic intuitions induced only by unusual offline tasks.

There are theoretical grounds for proposing such experiments. For example, in Barsalou's (1999) perceptual symbols systems, a simulator "controls attention across the simulation" (p.604). If our experiments have successfully tapped the spatial element of such simulators, concepts, or image schemas, then we would expect linguistic processing to modulate spatial attention in some manner. For example, if it is the case that the representation of certain words have a spatial element with some degree of verticality, then perhaps priming subjects with a vertical image schema such as those used in Experiment 1 would facilitate a lexical decision task for words such as 'respect' and 'succeed'.

In addition to more standard psycholinguistic paradigms, we hope to use eye movement data to investigate the spatial component of linguistic processing. It has been well demonstrated that mental imagery and mental models exhibit properties of an analog spatial layout (eg Denis & Cocude, 1992; Bower & Morrow, 1990). Work in our laboratory has demonstrated that this spatial component is evidenced in subjects' eye movements. When passively listening to a scene description and staring at a blank wall, subjects tend to make eye movements that correspond to the direction of the events described (Spivey, Tyler, Richardson & Young, 2000).

Similarly, Kaden, Wapner and Werner (1955) showed that visually perceived eye level is influenced by the spatial components of words. Subjects sat in a dark room and saw luminescent words at their objective eye level. Subjects then had the words moved up or down, until they were at the subjective eye level. Words with an upward connotation ('climbing', 'raising') had to be placed lower to be perceived as being at eye level, whereas words with a downward component ('falling', 'plunging') had to be placed above the objective eye level.

We hope that these image schema norms will allow us to measure spatial effects at a finer grain of representation than previous eye movement studies. Given that corresponding eye movements are made during an explicitly spatial description, perhaps they will also be made during an implicitly spatial description (due to metaphorical connections to image schemas). For example, we might find a bias towards vertical eye movements when listening to a description of John's respect for Mary.

Many of these findings are, or will be, surprising. If the purpose of communication is to *guide attention* and *coordinate action*, one should expect visuo-spatial information to play a causal role in linguistic processing. Future work will determine how strong a role.

Acknowledgements

We would like to thank Jamie Cheung for her assistance with data collection, and Dick Neisser for valuable discussions. Supported by a Sloan Foundation Fellowship in Neuroscience to MJS, and a Sage Fellowship to DCR.

References

Barsalou, L. W., (1999). Perceptual symbol systems. Behavioral and Brain Sciences,, 22(4) 577-660.

Boroditsky, L., (1999).First-language thinking for second language understanding: Mandarin and English speakers' conception of time. *Proceedings of theTwenty-first Annual Meeting of the Cognitive Science Society*, Erlbaum: Mawhah, NJ

Boroditsky, L., (2000) Metaphoric structuring: understanding time through spatial metaphors *Cognition*, *7,* 1-28.

Bower, G.H., & Morrow, D.G. (1990). Mental models in narrative comprehension. *Science, 247,* 44-48.

Carlson-Radvansky, L. A., Covey, E S., Lattanzi, K. M., (1999). What effects on "where": Functional influences on spatial relations. *Psychological Science, 10(6),* 516-521.

Chatterjee, A., (2001). Language and space: some interactions. *Trends in Cognitive Sciences, 5(2),* 55-61.

Denis, M., Cocude, M., (1992). Structural properties of visual images constructed from poorly or well-structured verbal descriptions *Memory and Cognition 20(5).*497-506

Gibbs, R. W., (1996). Why many concepts are metaphorical *Cognition, 61*, 309-319.

Gibbs, R. W., Strom, L. K., Spivey-Knowlton, M. J., (1997). Conceptual metaphors in mental imagery for proverbs. *Journal of Mental Imagery*, *21(3-4),* 83-109.

Hayward, W.G., Tarr, M.J., (1995). Spatial language and spatial representation. *Cognition*, *55(1),* 39-84.

Kaden, S.E., Wapner, S., & Werner, H., (1955) Studies in physiongomic perception: II. Effect of directional dynamics of pictured objects and of words on the position of the apparent horizon. *Journal of Psychology, 39,* 61-70.

Lakoff. G., (1987). *Women, Fire and dangerous things*. The University of Chicago Press: Chicago.

Landau, B., Jackendoff, R., (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16,* 217-265.

Langacker. R.W., (1987). An introduction to cognitive grammar. *Cognitive Science, 10(1),* 1-40.

Scheerer, M., & Lyons, J., (1957) Line drawings and matching responses to words. *Journal of Personality, 25*, 251-273.

Schober, M.F., (1995). Speakers, addressees, and frames of reference: whose effort is minimized in conversations about locations? *Discourse Processes*, *20,*219-247.

Spivey, M.J, Tyler, M., Richardson, D.C. & Young, E., (2000). Eye movements during comprehension of spoken scene descriptions. *Proceedings of the Twenty-second Annual Meeting of the Cognitive Science Society*, Erlbaum: Mawhah, NJ

Spivey-Knowlton, M.J., Tanenhaus, M., Eberhard, K., & Sedivy, J., (1998). Integration of visuospatial and linguistic information: language comprehension in real time and real space. In Oliver & Gapp (eds) *Representation and processing of Spatial Expressions*. LEA: London.

Talmy. L., (1983). How language structures space. In Pick and Acredolo (eds), *Spatial orintation: theory, resarch and application.* Plenum Press: New York.